



TITLE:

FMO-MO 法による大規模分子軌道計算(計算科学の基盤技術とその発展)

AUTHOR(S):

稲富, 雄一; 梅田, 宏明; 渡邊, 寿雄; 長嶋, 雲兵; 櫻井, 鉄也

CITATION:

稲富, 雄一 ...[et al]. FMO-MO 法による大規模分子軌道計算(計算科学の基盤技術とその発展). 数理解析研究所講究録 2006, 1505: 134-144

ISSUE DATE:

2006-07

URL:

<http://hdl.handle.net/2433/58511>

RIGHT:

FMO-MO 法による大規模分子軌道計算

九州大学情報基盤センター 稲富 雄一 (Yuichi Inadomi)

Computing & Communications Center,

Kyushu University

産業技術総合研究所計算科学研究部門、科学技術振興機構戦略的創造研究推進事業

梅田 宏明 (Hiroaki Umeda)

渡邊 寿雄 (Toshio Watanabe)

長嶋 雲兵 (Umpei Nagashima)

Research Institute for Computational Sciences (RICS),

National Institute of Advanced Industrial Science and Technology (AIST)

Core Research for Evolutional Science and Technology (CREST),

Japan Science and Technology Agency (JST)

筑波大学大学院システム情報工学研究科 櫻井 鉄也 (Tetsuya Sakurai)

Institute of Information Sciences and Electronics,

University of Tsukuba

1. はじめに

量子力学に基づいた第一原理計算手法の1つである分子軌道 (Molecular Orbital, MO) 法は、原子・分子の中に存在する電子の様子 (電子状態) を求めるための計算方法である。様々な化学現象には、必ず、電子状態の変化を伴っているため、MO法による計算を行うことで、化学現象を理論的に解明することができる。MO法による計算では、分子中の電子がどのように分布しているかを表す関数である、分子軌道、と呼ばれる関数が数多く得られる。そのうち、特に、フロンティア軌道と呼ばれる2つの軌道と、その軌道に含まれる電子が持つエネルギー (軌道エネルギー) が、様々な化学反応のメカニズムを解析する際に重要な役割を果たすことがノーベル賞 (化学賞) を受賞された故福井謙一博士によって示されている[1]。フロンティア軌道を求めることで、分子のどの部分で反応が起こりやすいか、どのような分子と化学反応を起こしやすいか、といった知見が得られる。近年、研究が盛んに行われている DNA やタンパク質などの生体分子が関与する一連の反応も、化学反応の一種であるため、巨大分子である生体分子の反応部位 (活性部位) を求めたり、生体分子との薬物との反応性 (薬理活性) を推定したりする用途にも、MO 法による理論計算を応用することができる。しかし、Hartree-Fock MO 法 (HF 法) [2]をはじめとする従来の MO 計算法では、分子サイズの4乗に比例する数をもつ2電子積分の計算と、それらを基に作られる大規模密行列 (Fock 行列) に対する一般化固有値問題求値を、10~100回程度繰

り返す必要があるため、大型高速計算機を用いた場合でさえも、小型生体分子の計算に数日～数ヶ月かかる。たとえば、創薬の分野では、実際の化学物質を用いた実験を行う前に、目的に適合しそうな候補化合物を絞り込む、スクリーニング、と呼ばれる操作が必要となるが、こうした分野に MO 法を適用するためには、数多くの化合物に対する計算を高速に行う必要がある。したがって、より大きな分子に対する計算や、多数回の分子軌道計算を行うためには、従来の分子軌道法で必要なこの膨大な計算時間を劇的に短縮することが必要不可欠である。

北浦らが提案したフラグメント分子軌道 (Fragment MO, FMO) 法[3-6]は、高速に分子の電子状態を求めるために開発された計算手法の 1 つである。FMO 法では、巨大な分子をいくつかの小さなフラグメントに分割して、各フラグメント (monomer)、およびフラグメントペア (dimer) に対する小規模な MO 計算を行うことで、分子全体の電子状態を求める。FMO 法は、HF 法と同等の計算精度を維持しつつ、計算量が非常に少なく、さらに並列処理にも適しているため、数 10 分～数日程度で、1000 原子を超える大規模分子のエネルギーや電子密度分布などを求めることができる。一方で、FMO 法そのままでは、分子全体の分子軌道を求めることができない、という欠点を持つ。分子の反応性を議論するためには、フロンティア軌道を求めることが必要であるため、FMO 法の利点を生かして分子全体の分子軌道を計算するための計算方法が求められていた。筆者らが開発した FMO 法に基づく分子軌道 (FMO-MO) 法[7]は、FMO 法の結果を利用して、巨大分子全体の分子軌道、ならびに軌道エネルギーを精度よく求めるための方法である。本稿では、従来の分子軌道法、FMO 法、および、FMO-MO 法について、適用事例などを交えながら解説を行う。

2. 理論

ここでは、代表的な非経験的 MO 法の 1 つである HF 法の概略と HF 法を大規模分子に適用する場合の問題点を述べ、その問題を解決するための HF 法を基にした FMO 法および FMO-MO 法の説明を行う。

2. 1 Hartree-Fock 分子軌道法

Hartree-Fock(HF)法は非経験的 MO 法としては比較的計算量が少ないため、大きな分子の MO 計算などに広く利用されている。HF 法において、分子軌道 $\{\psi_a(\mathbf{r})\}$ は「基底関数」と呼ばれる既知の関数 $\{\phi_i(\mathbf{r})\}$ の線形結合で表される。

$$\psi_a(\mathbf{r}) = \sum_i^N (c^a)_i \phi_i(\mathbf{r})$$

ここで N は計算に用いられる基底関数の数であり、分子のサイズや要求される分子軌道の精度に依存する。HF 法とは、以下に示す Hartree-Fock-Roothaan (HFR) 方程式という行列方程式を解くことによって MO の展開係数 (MO 係数ベクトル, $\{c^a\}$) を求める方法である。

$$\mathbf{FC} = \mathbf{SC}\epsilon$$

ただし、

$$\begin{aligned} \mathbf{F} &= \mathbf{H}^{\text{core}} + \mathbf{G}, & H_{ij}^{\text{core}} &= \int d\mathbf{r} \phi_i(\mathbf{r}) \hat{h} \phi_j(\mathbf{r}) \\ G_{ij} &= \sum_{k,l} D_{kl} \left\{ (ij|kl) - \frac{1}{2} (il|kj) \right\}, & (ij|kl) &= \int d\mathbf{r}_1 \int d\mathbf{r}_2 \phi_i(\mathbf{r}_1) \phi_j(\mathbf{r}_1) \frac{1}{|\mathbf{r}_1 - \mathbf{r}_2|} \phi_k(\mathbf{r}_2) \phi_l(\mathbf{r}_2) \\ S_{ij} &= \int d\mathbf{r} \phi_i(\mathbf{r}) \phi_j(\mathbf{r}), & D_{ij} &= 2 \sum_a^{n/2} C_{ia} C_{ja}, & E &= \frac{1}{2} \text{Tr}(\mathbf{D}(\mathbf{H}^{\text{core}} + \mathbf{F})) \end{aligned}$$

ここで、行列 $\mathbf{F}, \mathbf{H}^{\text{core}}, \mathbf{G}, \mathbf{S}, \mathbf{D}$ は、それぞれ Fock 行列、核・1 電子ハミルトニアン行列、2 電子ハミルトニアン行列、重なり行列、および (電子) 密度行列と呼ばれる行列であり、 $N \times N$ のサイズを持つ。このうち、重なり行列 \mathbf{S} は正定値対称行列であり、その他は対称行列である。 E は電子エネルギー、 \hat{h} は 1 電子演算子で、 n は電子数を表している。また、 $\mathbf{C} \equiv (\mathbf{c}^m, \mathbf{c}^{m+1}, \dots, \mathbf{c}^p)$ ($1 \leq m \leq p \leq N$) は MO 係数ベクトルの組で構成される MO 係数行列である。HF 法では、HFR 方程式に Fock 行列と重なり行列を与えて一般化固有値問題を解くことになるのだが、MO 係数ベクトルを求めるために以下のような困難が存在している。

- (1) HFR 方程式で与えるべき Fock 行列が、密度行列を通して方程式の解である MO 係数行列に依存している。したがって HFR 方程式は非線形方程式であり、反復解法 (Self-Consistent Field(SCF)法) で解く必要がある。その繰返し回数は分子の大きさや種類にも依存するが、通常 10~100 回ほどである。
- (2) Fock 行列作成には、2 電子積分 ($(ij|kl)$) を使って求める 2 電子ハミルトニアン行列の計算が必要であるが、この積分の計算量が $O(N^4)$ となる。また、2 電子積分の数が外部記憶に保存することも困難なほど多いため、通常、反復のたびに計算する必要がある。
- (3) 最終的にフロンティア軌道 (に対応する固有値、固有ベクトル) しか必要でない場合でも、反復計算中は密度行列計算のために、最低、電子数の半分の固有ベクトルを求めなくてはならない。そのベクトルの数は、タンパク質などの巨大分子では、数万~10 万となる。

タンパク質などの大規模分子では基底関数の数が数万~数 10 万に及ぶため、HF 法を用いた MO 計算は、計算コストが非常に大きく困難である。

2. 2 フラグメント分子軌道 (FMO) 法

大規模分子に対する HF 法の最大の問題は 2 電子積分の計算量が分子サイズ N の 4 乗 $O(N^4)$ となることであった。これに対し FMO 法では、分子を 20~30 個程度の原子で構成される小さなフラグメントに分割して考えることにより、高速計算を可能としている。分割したフラグメントとフラグメントペアに対して HF 法を適用することで、各フラグメントならびにフラグメントペアに対するエネルギーや密度行列が求められる。FMO 法における分子全体のエネルギー (E^{FMO}) や密度行列 (\mathbf{D}^{FMO}) は、これらにより次のように与えられる。

$$E^{\text{FMO}} = \sum_{I>J}^{N_f} E_{IJ} - (N_f - 2) \sum_I^{N_f} E_I, \quad \mathbf{D}^{\text{FMO}} = \sum_{I>J}^{N_f} \mathbf{D}_{IJ} - (N_f - 2) \sum_I^{N_f} \mathbf{D}_I$$

ここで、 $\{E_I\}$ 、 $\{E_{IJ}\}$ ならびに $\{\mathbf{D}_I\}$ 、 $\{\mathbf{D}_{IJ}\}$ は、それぞれフラグメント I およびフラグメントペ

ア IJ のエネルギーならびに電子密度行列を表す。

各フラグメントやフラグメントペアに対する MO 計算は独立に行うことができるため、FMO 法では並列処理による高速化を効率よく行うことができる[5]。また分子全体を一度に扱うことがないため計算量は $O(N^2)$ になり、HF 法の $O(N^4)$ に比べて大幅に減少する。この結果、FMO 法では大規模分子のエネルギー、ならびに電子密度分布を、HF 法に比べて非常に短時間に、しかも精度よく求めることが可能である[4,5]。その一方で、各フラグメント、フラグメントペアの軌道エネルギー（固有値）や分子軌道（固有ベクトル）の計算は行うが、分子全体の軌道を求めることができない欠点を持っている。このため、タンパク質などの大規模分子のフロンティア軌道を求めるためには FMO 法を拡張する必要がある。

2. 3 FMO-MO 法

FMO 法では、分子全体の電子密度行列を得ることはできるが、分子軌道を求めることができなかった。FMO 法で計算できる分子のエネルギーや電荷分布なども反応メカニズムを理論的に解明するための有効な情報ではあるが、活性部位などの議論のためには、やはり分子軌道を求めることが望まれる。フラグメントの軌道で分子全体の軌道の代用をすることも提案されているが、比較的小さな分子でもフラグメントの軌道エネルギーの準位が HF 法の結果と違うことがあるなど問題も多い [8]。そこで著者らは、FMO 法で得られる分子全体の密度行列を基に、分子全体に対する Fock 行列と重なり行列を一度だけ計算し一般化固有値問題を解くことで、分子全体における分子軌道ならびに軌道エネルギーを求める計算方法を提案した。こうして得られた分子軌道を FMO 法における分子軌道 (FMO-MO) と呼び、これまでに FMO-MO が HF 法で得られる分子軌道とほぼ一致することが示されている[7]。表 1 に従来 MO 法と FMO-MO 法との相違点を示す。FMO-MO 計算では非線型方程式 (HFR 方程式) に対する近似を行っているため、それに基づいた HF 法で必要となる反復処理 (SCF 計算) を行わなくて済むため、分子全体に対する Fock 行列作成 (2 電子積分計算) をただ一度だけ行うだけでよく、高速な計算が可能である。またフロンティア軌道に対応するごく一部の固有値、固有ベクトルのみを計算すればよいことも FMO-MO 計算の特徴の一つである。

表 1：従来 MO 法と FMO-MO 法との比較

		従来 MO 法	FMO-MO 法
反復処理		必要(10~100 回)	不必要
2 電子積分計算回数		複数 (10~100) 回	1 回
一般化 固有値問題	求める固有値・固有ベクトルの数	数千~数万個	フロンティア軌道付近の数個~数十個
	解く回数	複数 (10~100) 回	1 回

2. 4 一般化固有値問題求値について — 櫻井-杉浦法 —

FMO-MO 計算では、大規模 Fock 行列と重なり行列を用いた一般化固有値問題を解く必要がある。表 1 に示したとおり、FMO-MO 法では、従来 MO 法とは異なり、フロンティア軌道付近の

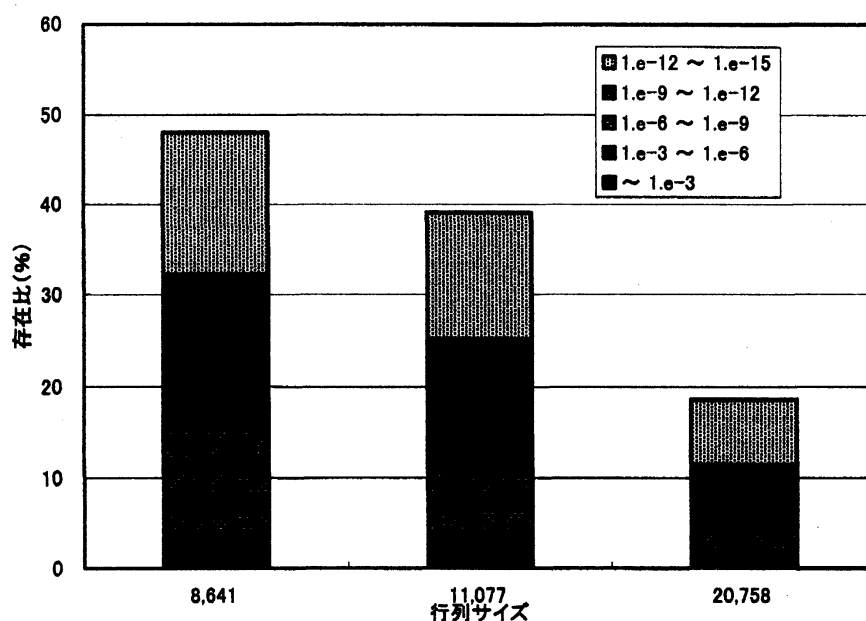


図 1 : Fock 行列の行列要素の大きさの分布

分子軌道，軌道エネルギーに対応する数個～数十個の固有ベクトル，固有値だけが必要であるため，そのような特性を生かした固有値解法が求められる。また，Fock 行列は比較的密な行列であるため，大規模疎行列に対する固有値解法の適用が困難である。図 1 に，異なるサイズを持つ分子に対する Fock 行列の要素の大きさの分布状況を示す。横軸は行列のサイズ，縦軸は与えられた数値範囲の大きさを持つ行列要素の存在比を示している。行列サイズが 8,641x8,641，11,077x11,077，20,758x20,758 と大きくなるにしたがって，大きな要素を持つ行列要素の割合は減少してはいるが，20,000x20,000 を超えるような比較的大規模な Fock 行列の場合でも， 10^{-9} 以上の大きさを持つ行列要素の割合が 5% 強と大きく，比較的密な行列であることが分かる。したがって，Fock 行列を疎行列として扱って固有値問題を解くことは，好ましくない。

以上に述べた特徴を持つ FMO-MO 計算での一般化固有値問題の解法として，今回は，櫻井・杉浦法[9]を用いることにした。詳細は原論文を参考にして頂くことにして，ここでは概要のみを述べる。櫻井・杉浦法では，ある数値区間に存在する固有値，固有ベクトルを求めるための大規模な一般化固有値問題を，多数の連立方程式と小規模な固有値問題の求値に置き換える。各連立方程式は独立に解くことが可能であり，さらに，連立方程式解法も並列化が行いやすいため，櫻井・杉浦法を用いた固有値問題解法は大規模並列化が可能である。また，後述するが，FMO 計算の結果を用いて，分子全体のフロンティア軌道の軌道エネルギーに対応する固有値の存在範囲の推定を行うことが可能なため，FMO-MO 計算での固有値問題解法として，櫻井・杉浦法との相性は非常に良いと考えられる。

3. 計算

FMO-MO を求めるためには，(1) FMO 計算，(2) FMO 計算で求めた密度行列を基にした分子全体の Fock 行列作成，および，(3) Fock 行列に対する一般化固有値問題求値，の 3 つのステ

ップが必要となる。このうち、はじめのステップ (1), (2) は、ともに並列処理向きの計算であり、高い効率で並列計算をすることができる。また最後のステップ (3) も、櫻井-杉浦法を用いることで、効率の良い並列計算が可能である。

今回の計算は、産総研に導入されている AIST super cluster[10]を構成する 3 つの大型クラスタ計算機のうち、P32 クラスタを使用して行った。このクラスタは、dual Opteron (model 246) マシン 1000 台あまりを、高速ネットワークの Myrinet で繋いだ大型 PC クラスタシステムである。FMO 計算には、国立衛生研の中野らが開発した並列 FMO 計算プログラム ABINIT-MP[5]を用いた。FMO-MO 計算における Fock 行列作成には、筆者らが開発した 2 電子積分プログラム[11]を用いた。これら 2 つのプログラムは MPI によって並列化されている。また、大規模 Fock 行列に対する一般化固有値問題求値には、櫻井らによってコーディングされ、Grid RPC の参照実装である Ninf-G [12]で並列化されたオリジナルコードを用いた。また、比較のために行った従来 MO 法での分子軌道計算には、計算化学の分野で広く用いられている GAMESS[13]を使用した。このプログラムも Distributed Data Interface (DDI) という独自の並列化ライブラリを用いて並列計算できるようになっている。ソースプログラムのコンパイルには PGI コンパイラ[14]を使用し、また、MPI ライブラリには、高速なクラスタとしての利用を容易にするために P32 クラスタに導入されている SCore クラスタシステムソフトウェア[15]が提供するものを用いた。固有値問題求値に関して、櫻井-杉浦法と比較するための従来固有値解法のプログラムとしては、AMD 社が提供している数値演算ライブラリ ACML[16]に含まれている LAPACK[17]ルーチンを用いた。

4. 結果と考察

4. 1 リゾチーム分子に対する計算結果 — 計算精度と計算時間 —

タンパク質などの大規模分子の活性部位を、分子軌道計算を用いて正しく求めるためには、フロンティア軌道の形状、位置が、精度よく計算することが必要である。そこで、まず、脊椎動物全般の細胞に広く含まれている殺菌作用のあるタンパク質の一種であるリゾチーム (Lysozyme) を対象分子として、従来 MO 法と FMO-MO 法でフロンティア軌道の計算を行った。このタンパクは、129 アミノ酸残基、1,961 原子で構成されており、分子軌道法での計算でよく用いられている基底関数の 1 つである STO-3G 基底関数系[18,19]を用いた場合、この分子の分子軌道を展開するために用いられる関数の総数は、6,005 である。図 2 に、この 2 つの方法で求めたフロンティ

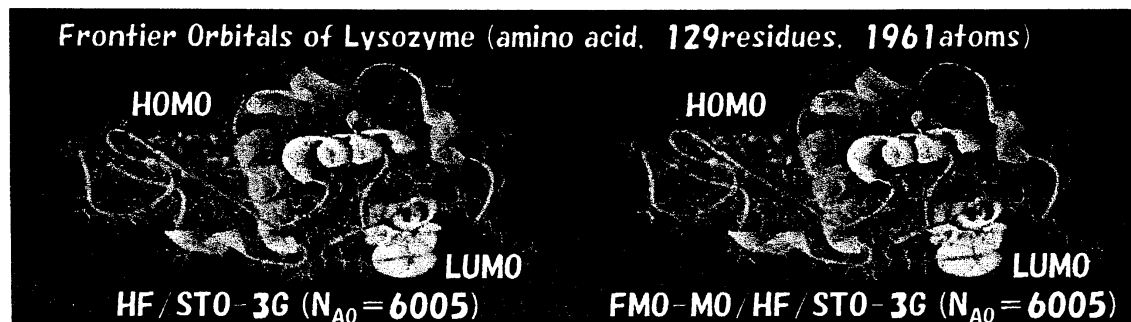


図 2 : 従来 MO 法(Hartree-Fock 法) (左図) と FMO-MO 法 (右図) で求めたリゾチーム分子のフロンティア軌道

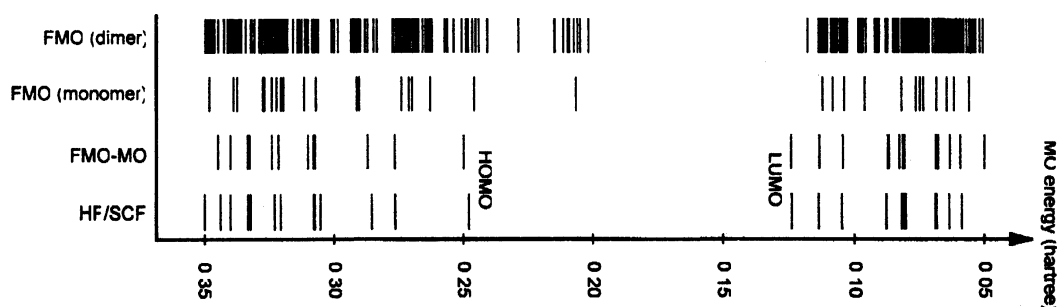


図3：フロンティア軌道付近の軌道エネルギー（固有値）分布

ア軌道を示す。図2の左図が、従来 MO 法で、また、右図が、FMO-MO 法で、それぞれ求めたフロンティア軌道の等値面を表わしている。また、各図にそれぞれ存在する2つの等値面のうち、左側は電子が詰まっているフロンティア軌道（Highest Occupied MO, HOMO）を、また、右側は電子の詰まっていないフロンティア軌道（Lowest Unoccupied MO, LUMO）を、それぞれ表わしている。この2つの図が区別できないほど似ていることから、FMO-MO 法は、フロンティア軌道の形状、および位置を、従来 MO 法と同等の精度で再現していることが分かる。

次に、フロンティア軌道付近の分子軌道に対する軌道エネルギー（固有値）の分布状況を図3に示す。この図には、従来 MO（HF/SCF）法、FMO-MO 法の結果に加えて、FMO 計算の途中で得られる各フラグメント（monomer）、フラグメントペア（dimer）の固有値分布も併記してある。従来 MO 法と、FMO-MO 法の結果を比較すると、この両者は非常に良く似た固有値分布を示していることが分かる。FMO-MO 法は、化学的に重要である2つのフロンティア軌道（HOMO, LUMO）の軌道エネルギー（固有値）の値、および、その差（HOMO-LUMO gap）を、従来 MO 法と同等の精度で算出している。先に述べた、フロンティア軌道の位置と形状、すなわち、固有ベクトルの情報だけではなく、その軌道エネルギー（固有値）も、高い精度で計算できている。したがって、FMO-MO 法は、フロンティア軌道付近の情報を求める目的への使用に適した計算手法であることが分かる。それに対して、FMO 法の計算で得られる各 monomer, dimer の軌道エネルギーの分布は、従来 MO 法や FMO-MO 法の結果から、かなりずれている。たとえば、HOMO の軌道エネルギーに相当すると思われる固有値が0.2 付近にあり、従来 MO 法の結果（-0.25）とは、約0.5も離れている。このずれは、化学的に見て許容できない大きな誤差であり、FMO 法の monomer, dimer の軌道エネルギー値の情報を、理論的な反応性推定に用いるのは好ましくない、と言える。

続いて、計算時間について述べる。表2に、リゾチーム計算にかかった、計算時間を示している。

表2：リゾチームのフロンティア軌道計算に要した時間（基底関数 STO-3G）

	計算時間（プロセッサ数）	
	HF/SCF/STO-3G	FMO-MO/HF/STO-3G
FMO 計算	one week (64 CPUs)	459 sec (544 CPUs)
Fock 行列作成		515 sec (512 CPUs)
一般化固有値問題		64 sec (16 CPUs)
全計算時間		1038 sec ≒ 17.3 min

ここで、GAMESS を用いた従来 MO 法の計算だけは、前出の P32 クラスタではなく、Quad Itanium2 (1.3GHz, メインメモリ 16GB/4CPU) マシン 120 台あまりを Myrinet で接続したクラスタ計算機 (M64 クラスタ) を用いている。これは、従来 MO 法の計算に用いたプログラム GAMESS が、FMO-MO 計算で用いた Fock 行列作成専用のプログラムよりも多くのメモリを消費すること、ならびに、P32 クラスタの主記憶容量が、M64 クラスタ計算機に比べて少ないためである。それはさておき、この結果から、従来 MO 法では、リゾチームのフロンティア軌道の計算におよそ一週間必要であったが、FMO-MO 法では、わずか 17 分強の計算時間で済んでいることが分かる。並列計算に用いたプロセッサ数に違いを考慮に入れても、FMO-MO 法での計算時間は、従来 MO 法に比べると、非常に短くなっている。これは、FMO 法による分子全体の密行列計算の計算量が従来 MO 法に比べると劇的に減少していることだけでなく、FMO-MO 計算に必要な 3 つの各計算ステップの並列化効率が非常に良いため、多くのプロセッサを用いた並列計算が行えることも、計算時間短縮の大きな要素である。

4. 2 モデル DNA に対する FMO-MO 計算

ここでは、DNA に対するテスト計算の結果を示す。DNA は生物の遺伝情報を保持する重要な生体分子であるばかりではなく、電気伝導性が高いことから、ナノレベルでの配線材料としての可能性も指摘されている興味深い分子の 1 つである。今回は、モデル DNA (TA-repeated sequence, 40 塩基対) 分子に対して FMO-MO 計算を行った。用いた基底関数は、リゾチームの場合と同じ STO-3G であり、原子数は 2,636、展開に用いた関数の数は 10,108 である。この規模の分子軌道計算は、従来 MO 法では大型計算機を用いても行うことが困難な計算規模であるが、FMO-MO 法では 500~600 プロセッサを用いた並列計算で、わずか 56 分で計算することが出来た。得られたモデル DNA の 2 つのフロンティア軌道を図 4 に示す。電子が詰まった軌道 (HOMO) を見ると、分子の中央部を中心に、広がった形をしていることが分かる。また、電子が入っていない軌道 (LUMO) が HOMO 付近に分布しているため、分子の熱運動に伴う揺らぎによって、HOMO に入っている電子が、容易に近接している LUMO に移動することが容易であると予想され、そのことが、DNA の電気伝導性のよさに関係していると思われる。

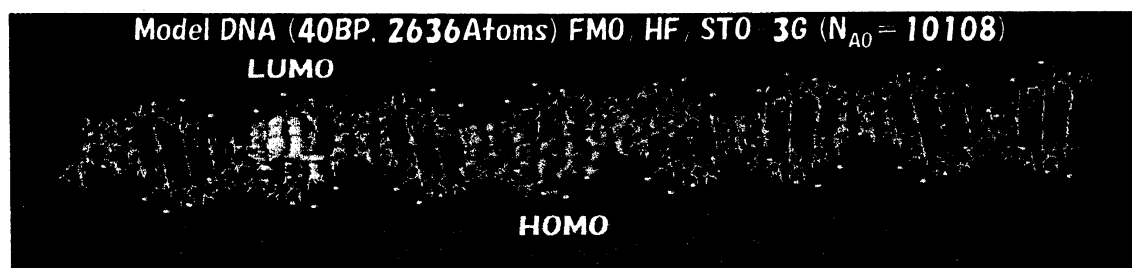


図 4 : モデル DNA のフロンティア軌道

4. 3 溶媒分子を含んだリゾチームに対する FMO-MO 計算

計算結果の最後に、溶媒 (水) 分子を露に含んだ、リゾチーム分子に対する分子軌道計算の結果

を示す。タンパク質などの生体分子が起こす様々な化学反応には、その周辺の水分子の存在が不可欠である。したがって、生体分子の反応メカニズムを理論的に正しく解明するためには、特に、生体分子と近接している水分子を含めた系に対する計算を行う必要がある。そこで、今回は、リゾチーム分子に、その電荷 (+9) を中和するための 9 個の陰イオン (Cl^-) とリゾチームから 10 Å 以内の距離にある 2,096 個の水分子を加えた、計 8,258 原子の巨大分子系に対する FMO-MO 計算を行った。その際に、STO-3G 基底関数形を用いたため、計算に用いた関数の総数は 20,758 である。図 5 に、フロンティア軌道付近の 2 つの分子軌道を示す。ここで大事なことは、この水を含んだ系に対する計算で得られた軌道と、先に示した水を含んでいない場合の結果が大きく異なる、という点である。次に、この計算に要した時間を表 3 に示した。まず、この計算全体を 2 日弱で計算できていることが分かる。また、FMO-MO の 3 つの計算ステップのうち、2 番目の Fock 行列作成が最も時間がかかっていることも分かる。しかし、Fock 行列計算は並列化効率が非常に高く、使用するプロセッサ数にほぼ比例して、速度向上が得られるため、より多くの計算機資源を用いることで、この大規模な分子軌道計算を数時間以内で行うことも可能である。また、最後のステップである一般化固有値問題求値も、並列化効率が上げにくい Cholesky 分解や Householder 変換, bisection 法などを組み合わせた一般的に分子軌道法で使われている固有値解法を用いた場合には、シングルスプロセッサ使用時に 15,000 秒以上かかっているが、櫻井・杉浦法では 128 プロセッサ使用時に、わずか 112 秒でフロンティア軌道付近の 7 つの固有値、固有ベクトルを計算できている。したがって、大規模分子の分子軌道計算を行うためには、従来 2 電子積分の計算時間に隠れてほとんど問題にされてこなかった固有値問題の計算時間を短縮する必要があることが分かった。また、フロンティア軌道に対応するごく少数の固有値、固有ベクトルしか

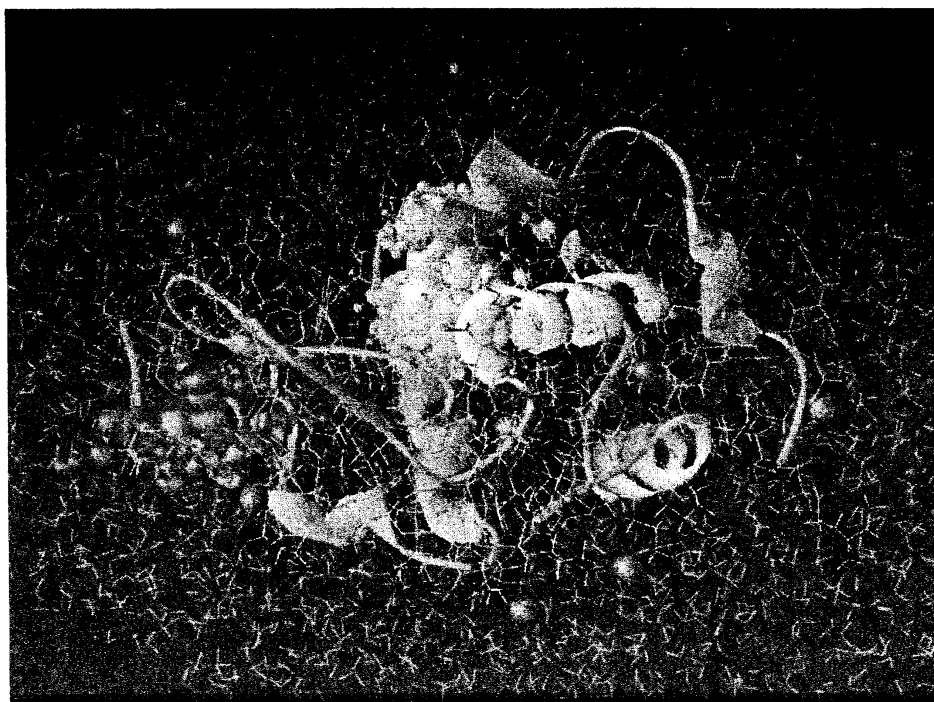


図 5 : 対陰イオンと周辺の水分子を含んだリゾチームの 2 つの分子軌道
(8,258 原子, 20,758 関数)

表 3 : 水分子等を含んだリゾチームに対する FMO-MO 計算にかかった時間

		プロセッサ数	計算時間 (秒)
FMO 計算		64	22,508
Fock 行列作成		64	147,051
一般化固有値問題	櫻井-杉浦法	128	112
	従来法*	1	15,011
全計算時間†		169,671 秒 (≒1.96 日)	

*: Cholesky 分解, Householder 変換, bisection 法, 逆反復法などを使用

†: 一般化固有値問題求値に櫻井-杉浦法を用いた場合

必要でないこと、ならびに、その目的の固有値の存在範囲が、FMO 法によって予測できることなどの、FMO-MO 法に特別の問題事情が、櫻井-杉浦法を大規模 Fock 行列に対して適用する際に、非常に有効であることが分かった。

5. まとめ

本稿では、フラグメント分子軌道法に基づいた分子軌道の計算手法である、FMO-MO 法、の紹介を行った。結果で示したとおり、FMO-MO 法を用いることで、従来の分子軌道法では困難であった、タンパク質などの巨大な分子に対するフロンティア軌道付近の分子軌道の計算を、高速に計算できることが分かった。また、数 1000 原子以上の大規模分子に対する分子軌道の計算を行う場合には、1000 原子未満の小・中規模分子の計算ではそれほど問題にならなかった一般化固有値問題求値を、より高速に行うことが必要であることが分かった。一般的な分子軌道計算における固有値問題解法として用いられている Householder 変換などを用いた、いわゆる直説法は比較的並列化が難しいとされている。しかし、今回用いた櫻井-杉浦法は、FMO-MO 法における固有値問題の特性に非常に適合した計算手法であり、かつ、高い並列化効率を持つため、大規模並列化による高速化も容易であり、2 万次を超える Fock 行列に対する固有値問題を、2 分弱という短時間に計算できた。したがって、FMO-MO 法を構成する 3 つの計算ステップすべてが、高い並列化効率で実行可能になった。FMO-MO 法を用いることで、今後開発されるであろう、次世代超並列計算機などにより、さらに短時間で分子軌道計算が可能になり、創薬などの分野への適用が容易になると思われる。

謝辞

この研究は、独立行政法人科学技術振興機構 (JST) が行う戦略的創造研究推進事業 (CREST) の研究領域「シミュレーション技術の革新と実用化基盤の構築」の研究プロジェクト「グリッド技術を用いた大規模分子シミュレーションプログラムの開発」の支援による。また、すべての計算は、独立行政法人産業技術総合研究所 (AIST) の大型クラスタ型計算機システム (AIST super cluster) を用いて行った。

参考文献

- [1] References in 米澤貞次郎, 永田親義, 加藤博史, 今村詮, 諸熊奎治, 三訂 量子化学入門 (上), 化学同人, 東京, 1990.
- [2] References in Szabo A. and Ostlund N. S., MODERN QUANTUM CHEMISTRY (Introduction to Advanced Electronic Structure Theory), Dover Publications Inc., Mineola, New York, 1996.
- [3] K. Kitaura, T. Sawai, T. Asada, T. Nakano, and M. Uebayashi, Chem. Phys. Lett. 312(1999)319.
- [4] K. Kitaura, E. Ikeo, T. Asada, T. Nakano, and M. Uebayashi, Chem. Phys. Lett. 313(1999)701.
- [5] T. Nakano, T. Kaminuma, T. Sato, Y. Akiyama, M. Uebayashi, and K. Kitaura, Chem. Phys. Lett. 318(2000)614.
- [6] K. Kitaura, S. Sugiki, T. Nakano, Y. Komeiji, M. Uebayashi, Chem. Phys. Lett., 336 (2001) 163.
- [7] Y. Inadomi, T. Nakano, K. Kitaura and U. Nagashima, Chem. Phys. Lett., 364 (2002) 139.
- [8] Sekino, H., Sengoku, Y., Sugiki, S. and Kurita, N., Chem. Phys. Lett., 378 (2003) 589.
- [9] T. Sakurai and H. Sugiura, J. Comput. Appl. Math., 159 (2003) 119.
- [10] AIST Super Cluster homepage, <http://unit.aist.go.jp/tacc/en/supercluster.html>
- [11] Y. Inadomi, T. Nakano, K. Kitaura, U. Nagashima, Increased efficiency of parallel calculations of fragment molecular orbitals by using fine-grained parallelization on a HITACHI SR8000 supercomputer, High-Performance Computing and Networking, 9th international conference, HPCN Europe 2001, Amsterdam, The Netherlands, June 2001, Proceedings
- [12] Ninf: A Global Computing Infrastructure, <http://ninf.apgrid.org/>
- [13] M.W.Schmidt, K.K.Baldrige, J.A.Boatz, S.T.Elbert, M.S.Gordon, J.H.Jensen, S.Koseki, N.Matsunaga, K.A.Nguyen, S.Su, T.L.Windus, M.Dupuis, J.A.Montgomery J. Comput. Chem., 14, 1347-1363(1993).
- [14] PGI Workstation 5.1: Copyright © 2003 The Portland Group™ Compiler Technology STMicroelectronics, Inc.
- [15] SCore homepage, <http://www.pccluster.org/index.html.en>
- [16] AMD Core Math Library (ACML) Version 2.1.0: Copyright © 2003, 2004 Advanced Micro Devices, Inc., Numerical Algorithms Group Ltd.
- [17] LAPACK -- Linear Algebra PACKage, URL:<http://www.netlib.org/lapack/>
- [18] W. J. Hehre, R. F. Stewart, and J. A. Pople, J. Chem. Phys., 1969, 51, 2657.
- [19] J. B. Collins, P. v. R. Schleyer, J. S. Binkley, and J. A. Pople, J. Chem. Phys., 1976, 64, 5142.